Neuropsychologia 198 (2024) 108881

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

The electrophysiology of lexical prediction of emoji and text

Benjamin Weissman [a,b,*], Neil Cohn [c], Darren Tanner [b,d]

[a] Department of Cognitive Science Rensselaer Polytechnic Institute 110 8th Street, Troy, NY, 12180, USA
[b] Department of Linguistics University of Illinois at Urbana-Champaign 707 S Mathews Ave, Urbana, IL, 61801, USA
[c] Department of Communication and Cognition Tilburg University PO Box 90153, 5000, LE Tilburg, the Netherlands
[d] AI For Good Lab Microsoft 1 Microsoft Way, Redmond, WA, USA

ARTICLE INFO

Keywords:
Emoji
Event-related potential (ERP)
Late frontal positivity (LFP)
Multimodality
N400
Prediction

ABSTRACT

As emoji often appear naturally alongside text in utterances, they provide a way to study how prediction unfolds in multimodal sentences in direct comparison to unimodal sentences. In this experiment, participants (N = 40) read sentences in which the sentence-final noun appeared in either word form or emoji form, a between-subjects manipulation. The experiment featured both high constraint sentences and low constraint sentences to examine how the lexical processing of emoji interacts with prediction processes in sentence comprehension. Two well-established ERP components linked to lexical processing and prediction – the N400 and the Late Frontal Positivity – are investigated for sentence-final words and emoji to assess whether, to what extent, and in what linguistic contexts emoji are processed like words. Results indicate that the expected effects, namely an N400 effect to an implausible lexical item compared to a plausible one and an LFP effect to an unexpected lexical item compared to an expected one, emerged for both words and emoji. This paper discusses the similarities and differences between the stimulus types and constraint conditions, contextualized within theories of linguistic prediction, ERP components, and a multimodal lexicon.

1. Introduction

The integration of text with emoji has quickly become a regular way that people communicate. Emoji meanings are integrated into the processing of sentences in real time (e.g., Barach et al., 2021; Beyersmann et al., 2023; Robus et al., 2020; Scheffler et al., 2022; Weissman and Tanner, 2018) and interact with grammar (Cohn et al., 2018, 2019), suggesting that these multimodal text-image interactions belong to a general communication system, as opposed to two separate ones (Cohn and Schilperoord, 2022). Here, we use this interaction between emoji and text to allow us to examine an area of language processing that has received much recent attention: prediction.

Ongoing work on language processing has debated whether, when, and the extent to which anticipation of upcoming information occurs at various levels of language (Carter et al., 2019; Heilbron et al., 2022), including lexical (e.g., Brothers et al., 2015; DeLong et al., 2014a), semantic (e.g., Federmeier et al., 2002; Luke and Christianson, 2016), morphological (e.g., Dillon et al., 2012; Luke and Christianson, 2015, cf. Lau et al., 2022), syntactic (e.g., Ferreira and Qiu, 2021; Lau et al., 2006), phonological (Connolly and Phillips, 1994; DeLong et al., 2005; van den Brink et al., 2001; cf. Poulton and Nieuwland, 2022) and orthographic structures (e.g., Laszlo and Federmeier, 2011). Much of this work has manipulated the predictability of a given word based on a prior sentence context; metrics of processing the target word (e.g., eye fixation probability, eye fixation duration, electrical brain activity) are taken as an index of the degree to which certain features of the word were anticipated or pre-activated. Some of these findings have recently been called into question (e.g., Nieuwland et al., 2018; Nieuwland, 2019; Nieuwland et al., 2020; Poulton and Nieuwland, 2022), situating feature anticipation and prediction at an epicenter of current sentence processing research. In the present study, we extend this line of work and investigate how sentence context modulates the lexical prediction of emoji. By examining issues of prediction with emoji, a set of familiar, conventionalized visual signs, we explore to what degree prediction operates at an abstract lexical level, and, in turn, how such interactions point towards an integrated multimodal communicative system. We adopt an electrophysiological approach with Event-Related Potentials (ERPs), focusing on the N400 and Late Frontal Positivity components as indices of predictive processes.

An important terminological clarification should be noted here. Ferreira and Lowder (2016: 218) describe prediction as the processing strategy that "make[s] use of contextual constraint to anticipate

* Corresponding author. Department of Cognitive Science Rensselaer Polytechnic Institute 110 8th Street, Troy, NY, 12180, USA.
E-mail addresses: bpweissman@gmail.com, weissb2@rpi.edu (B. Weissman).

https://doi.org/10.1016/j.neuropsychologia.2024.108881
Received 7 May 2023; Received in revised form 18 February 2024; Accepted 27 March 2024
Available online 4 April 2024
0028-3932/

upcoming input, leading to facilitated processing once the input is encountered." Thus, we might draw a distinction between anticipation, which refers to the pre-activation of upcoming features, and prediction, which describes the underlying processing strategy that may manifest as anticipation. A somewhat different distinction comes from Van Petten and Luka (2012) and Delong et al. (2014b), who use *prediction* to mean the preactivation of a specific linguistic item and *anticipation* to mean the preactivation of more abstract lexical features. While the two terms are sometimes used interchangeably, we will follow the latter distinction between the two terms for the purposes of this paper. Our primary focus is on the prediction of specific lexical items. By investigating lexical prediction in this way, we hope to shed light on processing meaning in a multimodal environment.

1.1. Multimodal processing

Recent work has argued that the verbal, gestural, and pictorial modalities can be integrated together into a single, multimodal cognitive architecture (Cohn, 2016; Cohn and Schilperoord, 2022). In this model, these different modalities link to a common conceptual structure of meaning, comprised of at least partially overlapping memory systems, as well as to combinatorial systems that structure their sequencing. Multimodality thus arises out of emergent coactivation of different structures in this broader architecture. Combining emoji with text is a novel instantiation, particularly for interactive digital communication, but such multimodality in general is intrinsic to the way humans communicate (e.g., Holler and Levinson, 2019; Zhang et al., 2021). In most text-emoji interactions, a common conceptual structure is shared between one modality that uses a grammatical system (text) and one that does not (emoji); these modalities interact to together express meaning.

Many multimodal interactions place one modality alongside another, such as gestures that temporally correspond with speech (e.g., Clark, 1996; Dimitrova et al., 2016; McNeill, 1992; Morett et al., 2020) or emoticons/emoji that are added to clauses of text (e.g., Barach et al., 2021; Christofalos et al., 2022; Filik et al., 2015; Holtgraves and Robinson, 2020; Pfeifer et al., 2022; Robus et al., 2020; Weissman and Tanner, 2018). However, *substitutive* interactions also occur prevalently, whereby units of one modality fill grammatical roles of another modality. In gesture processing, these are referred to as "language like" (McNeill, 1992), "component" (Clark, 1996), or "pro-speech" (Schlenker, 2019) gestures. Prior studies have suggested that images, too, are readily integrated into the semantics of a sentence (e.g., Ganis et al., 1996; Nigam et al., 1992), and more recent research suggests that multimodal substitutions (tested with gestures, animations, and pictures) can even lead the observer to generate semantic and pragmatic inferences as reliably as standard words (Hintz et al., 2023; Pérez et al., 2020; Schlenker, 2019; Tieu et al., 2019).

Emoji can also substitute for words in a sentence, like the familiar *I ❤ NY*. Recent work has observed that emoji replacing nouns and verbs can congruently be integrated into sentences (Ge and Herring, 2018). Cohn et al. (2018) found that emoji substitutions less congruent with their grammatical positions incur subsequent costs in processing downstream words. Cohn et al. (2019) elaborated on this account and found that emoji, while capable of being substituted in for words and contributing significant communicative content, are done so with a relatively simple (i.e., linear) grammatical structure and are more likely to replace nouns or adjectives than verbs or adverbs. Even noun substitutions, however, may incur a processing cost, as Paggio and Tse (2022) found significant differences in eye-tracking measures between nouns and corresponding emoji substitutions. The authors claim that "emoji used as word tokens are more difficult to integrate in the processing of the sentence than the word tokens themselves" (25), though it is also possible that these findings reflect a general modality-switching cost rather than anything about processing emoji themselves.

Research on electrophysiological responses to emoji substitutions has found that incongruous emoji can elicit certain neural patterns qualitatively similar to those elicited by incongruous words; however, these studies found differences between words and emoji in other neural signatures, interpreted as evidence for different mechanisms involved in word-only unimodal sentence processing and word + emoji multimodal sentence processing (Tang et al., 2020, 2021). It is worth noting, however, that the stimuli used in these experiments appear to feature relatively uncommon and unnatural emoji substitutions, as in: "Our project eventually succeeded, and I felt very ☺/happy/☹/unhappy." Face emoji, typically among the most commonly-used emoji, are usually utilized to highlight or modify tone, either accompanying text or as standalones, rather than as literal substitutions for emotion words, as illustrated by both qualitative (Gibson et al., 2018; Grosz et al., 2023; Yang and Liu, 2021; Yus, 2021) and corpus (e.g., Herring and Dainas, 2017; Li and Yang, 2018; Na'aman et al., 2017) analyses. The effects found in these studies may be due to the unnaturalness of the substitutions and may not extend to all (i.e., more naturalistic) emoji substitutions.

The current experiment aims to add to the ongoing investigation into the processing of emoji substitutions in real time by assessing well-established ERP signatures. Do they merely yield interpretable utterances or are they more readily integrated into sentences in real time?

1.2. N400

Psycholinguistics has a long tradition of using ERPs to study semantic processing, dating back to seminal works in the 1980s (e.g., Kutas and Hillyard, 1980, 1983, 1984). This line of work described the N400 – a negative-going brain potential, peaking around 400 ms post-stimulus and typically centro-parietal with a slight right hemisphere bias, sensitive to the processing of meaning (see Kutas and Federmeier, 2011; Nour Eddine et al., 2022 for thorough reviews). While all meaningful stimuli will elicit a negative-going component peaking around this time post-stimulus (the N400 component), contextually unexpected lexical items (such as "ankle" in "He poured maple syrup on his ankle.") generate a larger (more negative) N400 component compared to expected items ("He poured maple syrup on his pancakes."); this difference is called the N400 effect.

While there is still debate over the exact functional interpretation of the N400 (e.g., Brouwer et al., 2017; Kuperberg et al., 2020; Lau et al., 2009; Nieuwland et al., 2019; Nour Eddine et al., 2022; Rabovsky et al., 2018, and many more) and what process(es) the component reflects, there are numerous studies that detail what factors can modulate the component. The notion of constraint is particularly relevant to the present study. The cloze procedure (Taylor, 1953) asks participants to fill in a blank in a sentence and offer their best guess as to what the blanked-out word should be. Gathering many responses gives the cloze probability of a given word – the percentage of respondents who agree on a word to fill in the blank. Cloze probability has been shown to correlate with N400 amplitude (e.g., DeLong et al., 2014a; Kutas and Hillyard, 1984; Quante et al., 2018; Wlotko and Federmeier, 2012): the higher the cloze probability of a given word, the smaller the N400 component will be. Cloze probabilities offer a metric of contextual constraint – how strongly the preceding context leads to the expectation of a specific word. A highly constraining sentence will have a high cloze probability for a specific word ("The US flag features red, white and ___"), and a lowly constraining sentence will not have very high cloze probabilities for any word ("Her favorite animal is the ___"). This predictability manipulation is not limited to word processing, as cloze probability modulates N400 amplitudes in non-verbal narrative sequences as well (Coderre et al., 2020).

Previous research has also demonstrated N400 modulation in multimodal environments. Manfredi et al. (2017) presented visual narrative sequences with word substitutions: a comic strip in which one panel displayed a word instead of the comic panel. The words could be either descriptive "sound effect" words (e.g., "Punch!") or onomatopoeia words (e.g., "Pow!"). This study found modulations of the N400 effect whereby anomalous onomatopoeic and descriptive word panels elicited larger amplitudes than congruent counterparts. An N400 effect also emerges for gestures incongruent with semantics of visually-presented sentences (He et al., 2020).

Other work has specifically investigated written sentences with image completions, albeit novel, non-conventional ones. Nigam et al. (1992) presented participants with sentences in which the final word was replaced by a line drawing (i.e., a drawing of socks instead of the word "socks") and found that semantically anomalous pictures elicited an N400 effect identical in time course, amplitude, and scalp distribution to that elicited by words. In similar experiments, however, Ganis et al. (1996), Federmeier and Kutas (2001), and Pérez et al. (2020) all found that this cross-modal picture N400 had a slightly different scalp distribution than that to words – the picture N400 effect surfaced more frontally than the typical central-posterior word-evoked N400. A frontal negative component like this has also been found to surface in unimodal scenarios, including mismatching individual images (Holcomb and MacPherson, 1994), incongruous images in narrative sequences (Cohn et al., 2012), and incongruous objects in static visual scenes (Chen et al., 2022; Võ and Wolfe, 2013), as well as to concrete (vs. abstract) nouns (Lee and Federmeier, 2008).

1.3. Late frontal positivity

Prediction in sentence processing has also been associated with an anterior positivity that emerges around 600 ms post-stimulus, typically found at Fz and sometimes with a slight left-hemisphere bias, hereafter referred to as the Late Frontal Positivity (LFP). The documented circumstances leading to an LFP are encountering an unexpected yet plausible stimulus in a highly-constraining context (e.g., Brothers et al., 2015; Brothers et al., 2020; DeLong et al., 2014a; Federmeier et al., 2007; Quante et al., 2018; Thornhill and Van Petten, 2012). Encountering "banana" in "For the snowman's eyes the kids used two pieces of coal. For his nose they used a banana" would elicit both an N400 effect and an LFP as compared to the expected "For his nose they used a carrot." An unexpected and implausible item (e.g., "For his nose they used a groan") would elicit the N400 effect without an enhanced LFP (examples from DeLong et al., 2014a).

As the LFP is less documented than the N400, there has also been less discussion about the functional significance of this effect. Based on its appearance specifically in high-constraint sentences, Federmeier et al. (2007) suspect it may be related to a prediction mismatch, in the form of either recognizing the error or allocating resources to revise based on said error. (DeLong et al., 2014a) investigate the component more directly. Their finding that the component is elicited to unexpected plausible completions, but not to unexpected anomalous completions, argues against the prediction mismatch and in favor of the revision hypothesis; there is mis-prediction in both, but revision in only the former. They hypothesize it may instead reflect the participant's dependence on world knowledge in reconciling the plausible, yet unexpected, completion. Thornhill and Van Petten (2012) suggest that this component is "more related to lexical than conceptual factors" (p. 18), a suggestion supported in a follow-up investigation by Brothers et al. (2015). The latter group's findings led them to propose that lexical predictions (which are strong in high constraint environments) are not only pre-activated but also already somewhat integrated into the sentence; (DeLong et al., 2014a), too, suggest that highly expected completions may be pre-activated. Lastly, Brothers et al. (2020) and Kuperberg et al. (2020) suggest that this component indexes "successfully updating the comprehender's current situation model with new unpredicted information" (Kuperberg et al., 2020: 2). Encountering an unexpected word in a constraining context will necessitate a revision of that incremental and predictive situation model, reflected by the LFP.

Like the N400, the LFP has been observed in multimodal environments as well. In the Manfredi et al. (2017) study mentioned earlier, in which sound effect words (e.g., "Punch!") or onomatopoeia words (e.g., "Pow!") were substituted into visual narrative sequences, the sound effect words, elicited an LFP effect compared to the onomatopoeia substitutions. As the sound effect words have been shown to occur at a low frequency in comics (Pratha et al., 2016), this finding demonstrates that unexpected yet plausible stimuli can trigger this neural response in multimodal environments and provides further support for a predictive network that integrates stimuli across modalities. Pérez et al. (2020) recorded ERP responses to two types of pictures following an auditorily-presented context: "consistent" pictures, which match a specific inferred prediction licensed by the context (e.g., a picture of a polar bear after hearing "Many animals live in the North and South Pole. Because the ice in these regions is melting, some animals die in the water because they cannot find a patch of ice"), and "inconsistent" pictures (e.g., a picture of a penguin after that same context). The researchers only analyze the N400 effect in this paper, but visual inspection of the waveforms suggests there might be a similar plausible-yet-unexpected LFP effect as well (p. 138). LFP effects have also been found to unpredicted ("incoherent") images in image-only picture pairs and corresponding written sentences (Jouen et al., 2021), and unpredicted but sequentially reconcilable panels in visual narrative sequences (Cohn, 2021; Cohn and Foulsham, 2020; Cohn and Kutas, 2017).

1.4. Current study

The observed modulation of language processing by predictability raises an interesting question about image substitutions: might they, too, be modulated by the predictability of a sentence context? If so, that suggests lexical prediction is part of a more general meaning-processing system with coverage broader than just words; this can be taken as further evidence of an integrated lexical processing system that can pull from and indeed be modulated across different modalities. Such a finding would also indicate that participants are able to easily access form-meaning mappings of emoji during typical language comprehension tasks. As conventionality has been demonstrated to affect the processing of emoji out of context (Weissman et al., 2023), the present study seeks to explore how these conventional symbols, rather than novel pictures and line drawings, motivate multimodal lexical prediction.

Our experimental design examines how theories of lexical prediction bear on multimodal scenarios by presenting participants with high constraint sentences with expected/unexpected completions and low constraint sentences with plausible/implausible completions; these sentences were presented either unimodally, all text, and multimodally, with the sentence-final noun replaced by its emoji substitution. Our overarching question in this investigation is whether emoji, substituted into sentences for words, are processed in a manner similar to those words. In combining the theories of multimodality and lexical prediction outlined above, we expect ERP responses to emoji to closely mirror those to words, i.e., a centro-parietalN400 effect to unexpected emoji, regardless of constraint, and a late frontal positivityto unexpected but plausible emoji in highly constraining contexts.

2. Methods

2.1. Stimuli

The stimuli were 160 sentences – 80 high constraint and 80 low constraint. The last word of each sentence was a noun with a corresponding emoji with high naming agreement. The emoji utilized were concrete noun emoji, such as animals, foods, and objects. This choice was motivated by findings from the Cohn et al. (2019) production study showing that emoji substitutions of nouns are more common than of verbs or adjectives or adverbs, as well as other work on the widespread polysemy and ambiguity of facial expression emoji (e.g., Weissman et al., 2023) we sought to avoid in these stimuli. The form of the sentence-final noun was manipulated between subjects; each participant would always see either a word or an emoji. This setup navigates away from the potential for differing early attentional components and allows for analysis of the processes of meaning integration without elements of lower-level visual surprise.

To ensure the emoji used had high naming agreement across multiple participants, a norming study was carried out. 316 emoji were divided into two lists, and each list was presented as a survey to participants on Amazon Mechanical Turk. Participants (N = 35 per list) were asked simply to type in the name of each emoji that appeared in the survey. Participants who consistently gave joke answers (n = 3) were excluded and their responses were omitted. 160 nouns, all with emoji with high naming agreement were selected for the stimuli sentences; the average naming agreement for the emoji was 0.948 (SD = 0.05, range = 0.914–1). Participants in the main experiment who saw sentence-final nouns in emoji form were shown a naming survey with the 160 emoji after the ERP experiment; average naming agreement from these participants was 0.993 (SD = 0.024, range = 0.809–1).

160 sentence frames were developed such that there were 80 high constraint sentence frames and 80 low constraint sentence frames. A pair of experimental sentences within each constraint condition was created by replacing the sentence-final noun with an alternative noun, selected from the same set. This means each high constraint expected noun also appears exactly once as a high constraint unexpected noun in a different sentence in a different experimental list. The high constraint sentence frames had one expected completion (high cloze, high plausibility) and one unexpected completion (low cloze, high plausibility); the low constraint sentence frames had one plausible completion (low cloze, high plausibility) and one implausible completion (low cloze, low plausibility).

These items were normed to ensure the conditions behaved as expected in terms of expectancy and plausibility. Cloze probabilities were acquired by asking participants (total N = 255, originally divided into two lists, then followed by a third list testing revised items after the first pass) recruited through Amazon Mechanical Turk (hereafter *MTurk*) to provide the most likely completion to each sentence frame. The cloze results from each of the four conditions are presented in Table 1.

Plausibility ratings were acquired by presenting the sentence in full (with word completions, not emoji) to a new set of participants (N = 123) recruited from MTurk and asking them to judge the sentence for plausibility ("Does this sentence describe something that is possible?") on a 1–5 scale. The average plausibility scores for the four conditions are provided in Table 1 (right).

Stimuli were sorted into four experimental lists. Lists were counterbalanced such that each list contained 40 stimuli from each condition and each participant would see each sentence frame only once.

Example sentences are provided below in Fig. 1.

2.2. Participants

Participants in the experiment were 42 university students (10 male, 32 female, average age = 20.9). All participants were right-handed, monolingual English speakers, and reported no history of brain trauma, neurological impairment, or psychoactive medication. Two

Table 1
Results of norming for the four relevant conditions. Cloze range is 0–1, Plausibility range is 1–5.

Condition	Cloze expectancy of sentence frame	Cloze probability of completion	Mean plausibility rating (std. dev.)
High Constraint Expected	0.9	0.9	4.84 (0.25)
High Constraint Unexpected	0.9	0.001	4.23 (0.60)
Low Constraint Plausible	0.186	0.038	4.69 (0.45)
Low Constraint Implausible	0.186	0	1.52 (0.48)

High Constraint Expected	The chef carved the turkey with the knife The chef carved the turkey with the 🔪
High Constraint Unexpected	The chef carved the turkey with the scissors The chef carved the turkey with the ✂️
Low Constraint Plausible	Her favorite animal is the monkey Her favorite animal is the 🐒
Low Constraint Implausible	Her favorite animal is the avocado Her favorite animal is the 🥑

Fig. 1. Example stimuli in the two presentation formats and the four experimental conditions.

participants' data were omitted from final analysis due to high initial trial rejection percentages (>20%), making a final dataset of 40 participants – 20 who saw sentence-final words and 20 who saw sentence-final emoji. Participants provided informed consent according to University of Illinois IRB protocol and were compensated financially for their participation.

Previous research (Miller et al., 2016, 2017; Tigwell and Flatla, 2016) has investigated cross-platform differences between emoji, finding that the visual differences can lead to miscommunications and differing interpretations. Though these studies targeted face emoji, the non-face emoji have significant cross-platform visual differences as well. Since our object of study here is form-meaning mappings, we sought consistency and familiarity in emoji used in the experiment. For that reason, we used the iOS emoji and only recruited iOS users as participants. Participants were asked in a post-experiment survey to rate (on 1–10 scales) how familiar they are with emoji and how frequently they use emoji. The average rating was 9.25 (SD = 1.01) for familiarity and 7 (SD = 2.55) for personal usage.

2.3. Procedure

Participants were seated approximately 100 cm in front of a computer monitor and randomly assigned to one of eight experimental lists. Sentences were presented one word at a time, with each word remaining in the center of the screen for 300 ms followed by a 200 ms blank screen (500 ms SOA). The timing for the sentence-final emoji was the same. A basic comprehension question appeared after 35% of trials to ensure participants continued to pay attention throughout the experiment.

2.4. Data analysis

Continuous EEG was recorded from 28 tin scalp electrodes in standard and extended 10–20 locations in an elastic cap. Electrodes placed below the left eye and at the outer canthus of each eye (referenced offline in a bipolar montage) monitored eye movements. Channels were referenced during recording to the left mastoid and re-referenced offline to the algebraic mean of the right and left mastoids. Electrode impedances were held below 10 kΩ. EEG was amplified using a BrainAmpDC bioamplifier system (Brain Products GmbH) and digitized with a 1000-Hz sampling rate and an online analog 0.016–250 Hz bandpass filter. A 0.1–30 Hz bandpass (12 dB/octave roll-off) was applied to the continuous EEG offline.

All data processing was done in the EEGLAB and ERPLAB toolboxes in MATLAB (Lopez-Calderon and Luck, 2014). Epochs were time-locked to the sentence-final emoji or word, beginning 200 ms before and ending 1500 ms after presentation of the target. Independent Component Analysis (ICA) was run on the data to isolate blink and saccade components. If a blink or saccade component was detected based on visual inspection of the component topography and time course, this component was removed. After removal of these components, the average trial rejection was 4.6% overall (3.7% for participants who saw words and 5.5% for participants who saw emoji).

ERP components were defined as the mean voltage within an a priori time window. A time window of 300–500 ms was used for analyzing the N400 and a window of 600–1000 ms was used for analyzing the LFP. A 200 ms pre-stimulus baseline was used. Because the completion manipulation was different in high and low constraint sentences (expected vs. unexpected in high constraint, plausible vs implausible in low constraint), these two levels of constraint were analyzed separately. Within each level of constraint, there was a two-level factor of sentence completion (expectancy – expected vs. unexpected in high constraint – and plausibility – plausible vs. implausible in low constraint).

Midline electrode analyses were run with four levels of anteriority (Fz, Cz, Pz, Oz). Four regions of interest, each including five electrodes, were used for analysis of lateral electrodes, with two levels of anteriority (anterior, posterior) and two levels of hemisphere (left, right). However, across all analyses the results of the lateral electrode models generally mirror those of the midline electrode models; for the sake of simplicity in reporting and avoiding extraneous modelling, the current analyses involve the midline electrodes only; the anteriority approach allows for assessment of the expected effect topographies, namely a Pz-centered N400 and an Fz-centered LFP.

In the spirit of moving away from null hypothesis testing (Cumming, 2014), our statistical analysis here presents visual/numerical depictions of effect sizes rather than *p*-values; this approach follows similar efforts in ERP research such as Payne et al. (2015) and Rodríguez-Gómez et al. (2020). The linear mixed effect model parameter estimates with 95% confidence intervals are presented visually/numerically for interpretation of effect sizes. If the confidence interval of a parameter does not overlap with zero, that may be interpreted as roughly equivalent to "statistical significance," though attention should be paid to the size of the effect, represented continuously, rather than as a significant/not-significant binary.

Primary analyses were conducted with linear mixed effects models on trial-level data (see Kretzschmar and Alday (2023) for more on this method of analysis) with random intercepts and slopes by sentence completion for participant and item. Grand mean ERPs are used here for scalp topography maps and waveform plots, but all analyses are conducted on the trial-level data. Sum coding was used for contrasts in all models, which means parameter estimates should be interpreted as effect size for that factor when averaged across levels of the other factors. This setup eliminates the need for post-hoc testing, and main effects can be interpreted directly from model outputs (Brehm and Alday, 2022; Schad et al., 2020). The independent variables of expectancy (expected vs. unexpected), plausibility (plausible vs. implausible), and stimulus (word vs. emoji) all have two factors and are coded as (−1, 1). Electrode has four levels (Fz, Cz, Pz, Oz), and codes Oz as the reference level (−1); the parameter listed for Fz, for example, would estimate the extent to which amplitude at Fz differs from the average amplitude across all midline electrode sites.

Since our hypotheses only concern electrode and stimulus type inasmuch as they relate to the expectancy and plausibility conditions, only the electrode*condition and stimulus*condition interaction terms are presented in results; this achieves the same purpose as analyzing difference waves to compare responses between emoji and words. Full model outputs are available at the link in the supplemental materials.

3. Results

3.1. High constraint

Grand mean waveforms from sentence-final words and emoji in high constraint sentences are displayed in Fig. 2, with scalp maps in Fig. 3.

Fixed effect parameter estimates from the linear mixed effects models run for high constraint sentences in both time windows are presented in Fig. 4.

In the 300–500 ms time window, there is evidence for a reliable

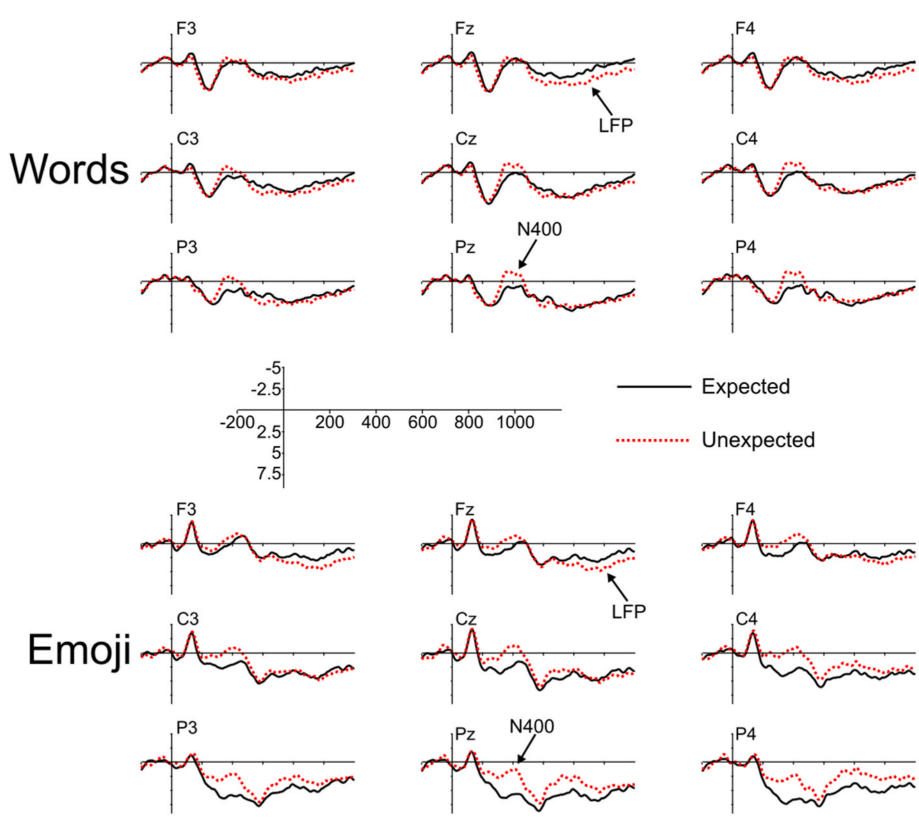

Fig. 2. Grand average ERP waveforms from nine representative electrodes for word and emoji completions to high constraint sentences. Negative is plotted up.

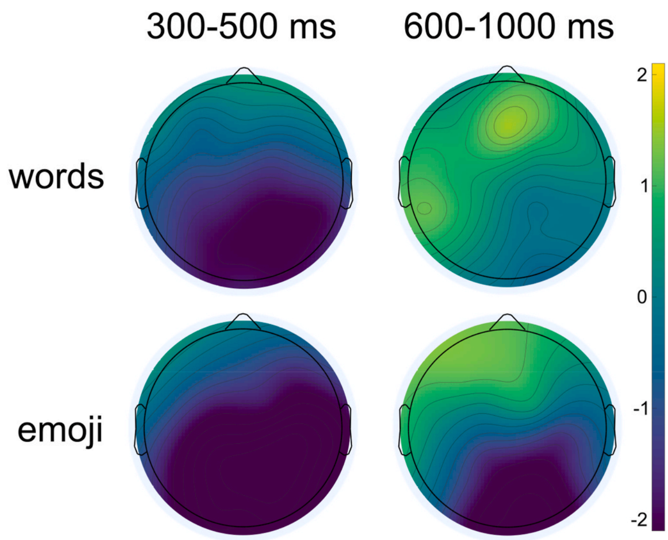

Fig. 3. Scalp maps derived from difference waves in the two relevant time windows from word and emoji completions to high constraint sentences.

overall effect of expectancy ($b = -1.00$; 95%CI = [−1.51 -0.49]). This negativity had a posterior scalp distribution, as evidenced by the unexpected*Pz interaction ($b = -0.42$; 95%CI = [−0.70 -0.13]. There is no evidence for a reliable interaction between expectancy and stimulus type during this time window, neither as modulated by electrode site nor overall.

In the 600–1000 ms time window, the expectancy effect is modulated by stimulus type, as evidenced by an increased negativity for emoji as compared to words for unexpected sentence completions, signaled by the unexpected*emoji interaction ($b = -0.41$; 95%CI = [−0.74 -0.07]). As in the earlier time window, anterior electrodes display a more positive mean amplitude than do posterior electrodes; per the visualizations in Figs. 2 and 3, this is interpreted as evidence of a Late Frontal Positivity emerging to unexpected completions, with statistical support provided by the reliable unexpected*Fz interaction ($b = 0.84$; 95%CI = [0.56 1.12]).

3.2. Low constraint

Grand mean waveforms from sentence-final words and emoji in low constraint sentences are displayed in Fig. 5, with scalp maps in Fig. 6.

Fixed effect parameter estimates from the linear mixed effects models run for low constraint sentences in both time windows are presented in Fig. 7.

In the 300–500 ms time window, there is an overall negativity for implausible sentence completions ($b = -0.46$; 95%CI = [−0.87 -0.05]; there is no evidence that this implausibility effect varies based on stimulus type nor electrode.

4. Discussion

We investigated the extent to which lexical predictive processing is modulated by context across modalities, assessing word versus emoji sentence completions. We compared ERPs between expected and unexpected completions under high constraint and plausible and implausible completions under low constraint. We found that the effects under investigation, namely an N400 to unexpected nouns and an LFP to unexpected yet plausible nouns were reliably elicited by both words and emoji in both environments. In neither constraint condition was there any statistical evidence of a difference between emoji and words in the 300–500 ms time window; however, there were small stimulus-related differences in the 600–1000 ms time window. Taken altogether, results indicate that participants accessed and integrated lexical meaning from emoji in these multimodal sentences.

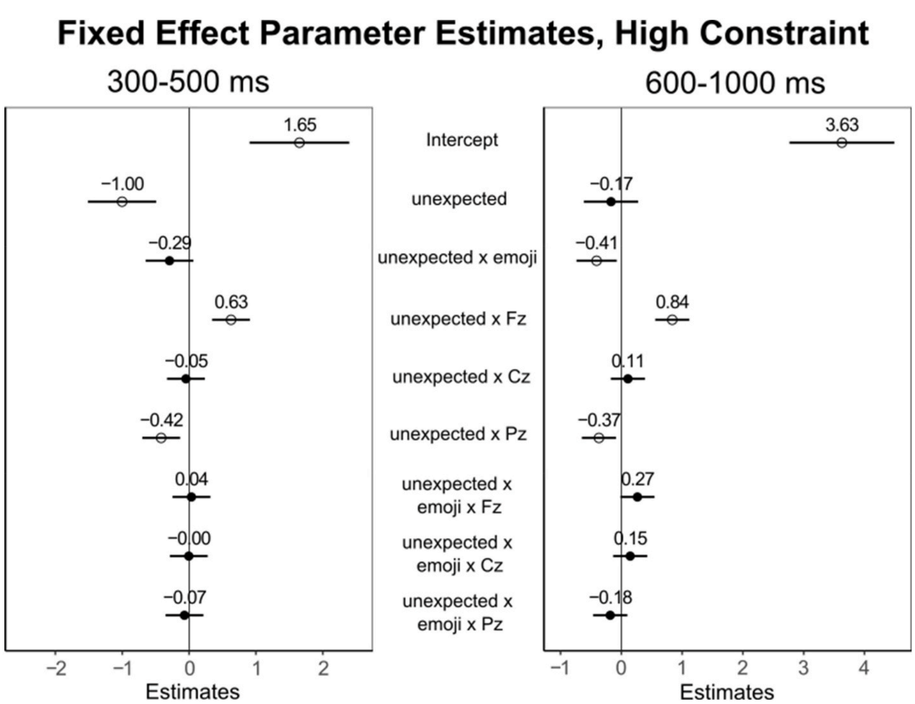

Fig. 4. Fixed effect parameter estimates from linear mixed effects models run (separately) on the two relevant time windows for high constraint sentences. Dots indicate parameter estimate, bars extend to 95% confidence intervals. Shaded dots appear for estimates with confidence intervals that include 0; unfilled dots appear for estimates with confidence intervals that do not include 0. Sum coding is used, with "expected," "word," and "Oz" used as reference levels for Condition, Stimulus, and Electrode, respectively.

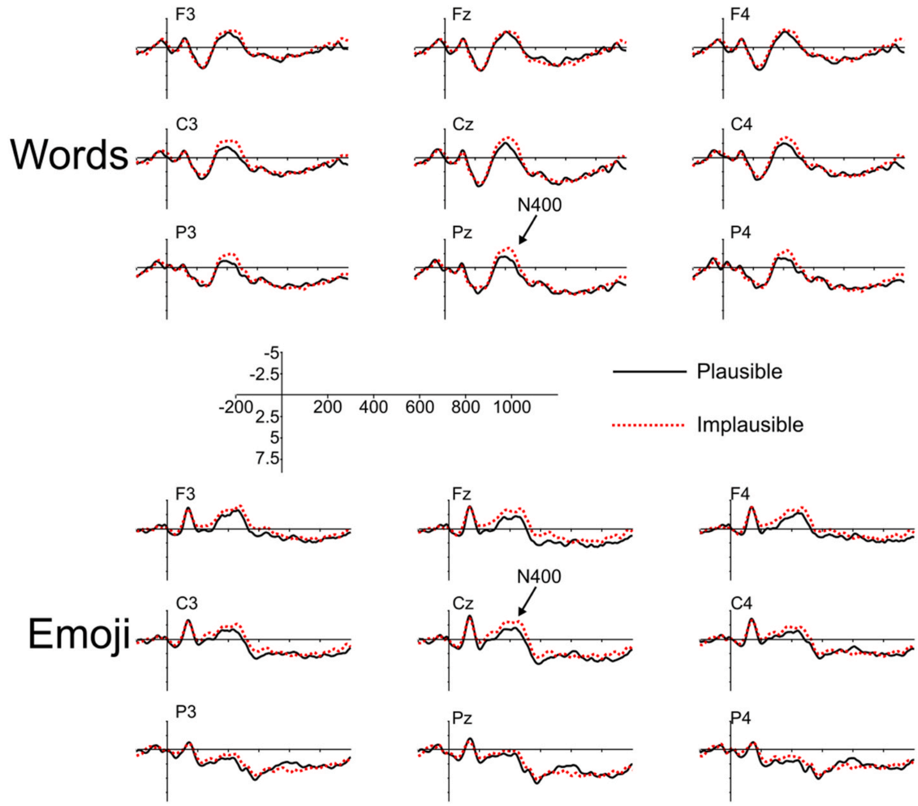

Fig. 5. Grand average ERP waveforms from nine representative electrodes for word and emoji completions to low constraint sentences. Negative is plotted up.

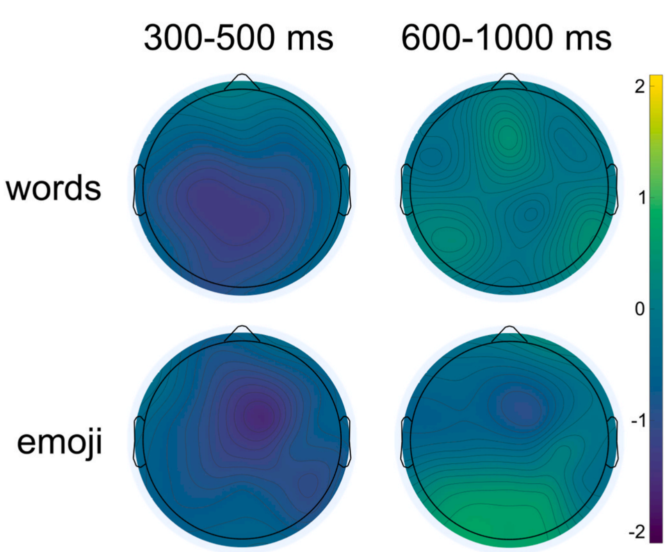

Fig. 6. Scalp maps derived from difference waves in the two relevant time windows (300–500 ms left, 600–1000 ms right) from word (top) and emoji (bottom) completions to low constraint sentences.

4.1. High constraint

In high constraint sentences in the 300–500 ms time window, an effect of expectancy, with a posterior distribution typical of the N400, appeared for both emoji and words. There was no evidence for any stimulus-related interactions with expectancy in this time window, suggesting similar N400-related processes for emoji and words. With flexibility based on preferred N400 interpretation, we can say that participants' lexical access was not affected by whether the lexical item appeared in word form or emoji form in these high constraint sentences, at least in a between-subjects setup in which the modality of the sentence-final noun itself is expected.

That an LFP effect was found to unpredicted yet plausible emoji suggests participants are also able to integrate emoji in a manner similar to words; both stimulus types evoked a classically-frontally-distributed positivity in the 600–1000 ms time window. Theories of this component (e.g., Brothers et al., 2020; Kuperberg et al., 2020) posit that it reflects an updating of the discourse model based on the unexpected lexical encounter. In this experiment, then, participants not only anticipated upcoming lexical content, regardless of modality, but also were able to rapidly access that lexical content and update their discourse model, again regardless of form. This interpretation is consistent with a multimodal conception of the lexicon (Cohn and Schilperoord, 2022), in which the lexicon includes not just items from the verbal modality (i.e., words) but meaning-making units from any modality (e.g., drawings, gestures, emoji). Under this light, the LFP component can be said to reflect a response that appears across modalities and their integrations. The continued exploration of any-modal situations that evoke this pattern contribute to the developing understanding of the process indexed therein.

It is worth noting, however, a stimulus-related difference in this time window, best characterized as an increased negativity to unexpected emoji. This negativity does not extend to frontal electrode sites and appears to be distinct from the aforementioned LFP, which does reliably emerge over those frontal sites, and instead resembles a "lingering" of the N400 component, best visualized at Pz/P4/C4 at the bottom of Fig. 2. Similar extended N400 effects have been found elsewhere in the literature in a wide range of cases, including but not limited to animacy violations (Vega-Mendoza et al., 2021), subject-verb agreement violations (Tanner, 2019), gender-mismatching articles (Fleur et al., 2020), inconsistent words in counterfactual sentences (Wang and Xu, 2022), violations of world knowledge in an L2 (Romero-Rivas et al., 2017), violations of world knowledge in an L1 (Leuthold et al., 2015), and

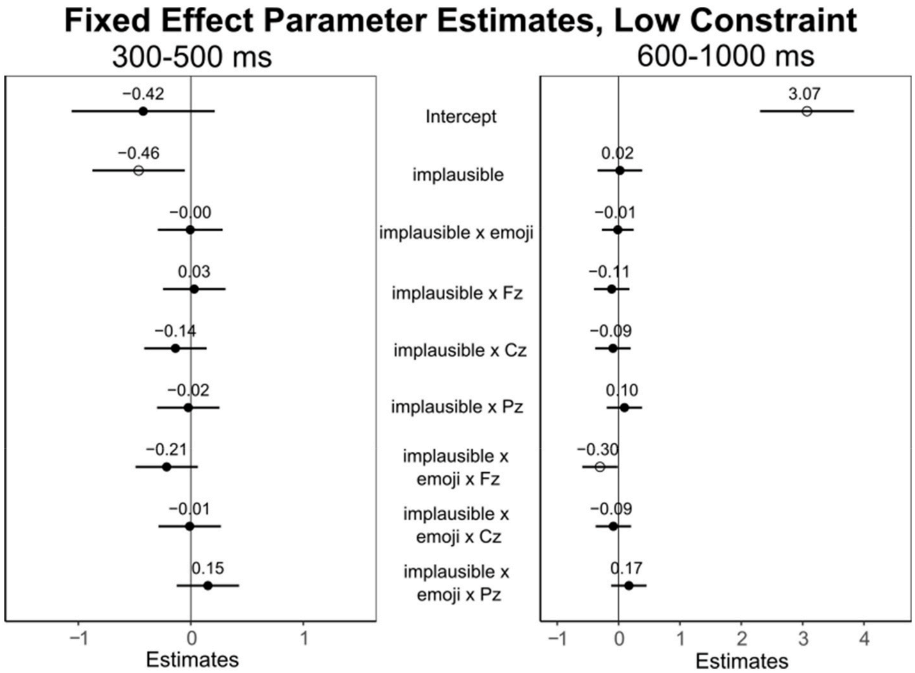

Fig. 7. Fixed effect parameter estimates from linear mixed effects models run (separately) on the two relevant time windows for low constraint sentences. Dots indicate parameter estimate, bars extend to 95% confidence intervals. Shaded dots appear for estimates with confidence intervals that include 0; unfilled dots appear for estimates with confidence intervals that do not include 0. Sum coding is used, with "expected," "word," and "Oz" used as reference levels for Condition, Stimulus, and Electrode, respectively.

various forms of sentence incongruity (e.g., Coulson et al., 2005; Meβmer et al., 2021). The temporal extension of the N400 has not been the subject of robust theoretical work; most often, authors will simply note that the N400 components they detected in the typical N400 time window also seem to extend beyond this time window, as we are doing here. We add this finding to the growing pile of extended N400 instances and further the call for theoretical explanations as to why this component sometimes, but not always, lingers.

Nieuwland et al. (2019) do offer an explanation of the effect's extension, positing that it could reflect "processing consequences of sentence plausibility and … continued efforts to integrate a word with its context" (7). At the very least, our current findings suggest broadening this claim from "word" to "lexical item," as emoji in this experiment elicited the extended effect. Per the Nieuwland et al. hypothesis, this could constitute evidence of an emoji-word processing difference, indexing extended integration difficulty of emoji but not of words.

Alternatively, there is another account of an extended post-N400 negativity, dubbed the N700 by West and Holcomb (2000). This N700 is intriguing, as it has been linked to increase word concreteness and mental imagery processes (e.g., Barber et al., 2013; Bechtold et al., 2023; Bechtold et al., 2018; West and Holcomb, 2000). These notions have ostensible connections to emoji processing, but the N700 effects described above all surface more over anterior sites, whereas the post-N400 negativity in this study is posterior. In addition to the topographical difference, the negativity in this study is less easily categorized as a distinct effect, as the expected-unexpected difference at Pz for emoji is present throughout the entire post-stimulus time course. Though not often highlighted, an effect resembling this has been found in other studies of the semantic processing of visual content: Cohn (2021) found a similarly posterior sustained negativity to visual narrative sequences in which the climactic panel is an inference-generating "action star" as compared to when that panel is an explicit depiction of the climactic event itself. Visual inspection of the waveforms in Pérez et al. (2020) suggests there may, too, be a post-N400 negativity to the picture completions in that study, though there is no statistical analysis of the later time window in the paper.

Given the relative underdevelopment of theories describing (a) exactly when and why the N400 effect extends and (b) processing mechanisms underlying the N700, we remain unable to draw strong conclusions at the present time. It is also the case, as Nieuwland et al. (2019) note, that there is likely significant neural complexity underlying detected signals across a larger time window like this, so these two accounts are not even mutually exclusive. Future work, especially within multimodal processing environments, may be able to begin establishing differentiating lines between these hypotheses.

That participants reliably accessed the meanings of the emoji, evidenced by the N400 and LFP effects to unexpected completions in high constraint sentence frames, comes as a natural consequence of the conventionality of these signs. It may be the case that the participants in this experiment, college-age iPhone users highly familiar with this set of signs, are particularly adept at this sort of processing. Repeated practice of accessing this form-meaning mapping may facilitate the process more so than seeing novel, unfamiliar line drawings, consistent with other findings demonstrating that conventionality can affect the lexical access of emoji out of context (Weissman et al., 2023). Participants in this experiment all self-reported extremely high familiarity and frequency-of-use of emoji to the point there is not enough variability to check for correlations between these measures and the relevant ERP components in the current data; future work with a differently-targeted participant recruitment should probe the extent to which an individual's familiarity with emoji and a given emoji's conventionality affect the neural correlates of lexical access. As this experiment utilized relatively monosemous, unambiguous emoji in highly constraining sentences, lexical access of these emoji is likely facilitated compared to in Tang et al. (2020). The emoji vs. word ERP differences in the 300–500 ms window from that study may be a consequence of the less conventionalized face emoji and the less highly-constraining sentences used in the experiment.

4.2. *Low constraint*

Implausible completions of low constraint sentences generated a significant negativity compared to plausible completions in the 300–500 ms time window for both emoji and words. As before, the lexical access-related process reflected by the N400 seems to activate regardless of stimulus modality. There is borderline evidence for a three-way interaction, observed within the implausible*emoji*Fz term ($b = -0.30$; 95% CI = [−0.59 -0.01]), interpretable as a weak sustained negativity to implausible emoji at Fz. This time window is not relevant for our hypotheses regarding the plausibility manipulation beyond the absence of any LFP effect; this coupled with the borderline evidence for the small effect leads us away from making any strong interpretations here. We will note, however, that this anterior negativity qualitatively resembles the aforementioned extended N400/potential N700 observed in the early window, once again.

Many of the low constraint stimuli in this experiment are simple sentences that essentially amount to categorization probes (e.g., "her favorite animal is the ฑ/") – these sentences may be prompting a more categorization-style form of processing than do other low constraint sentences. Further work could investigate the extent to which sentence constraint and explicit or implicit categorization processes guide people to process emoji differently in different contexts by way of a systematic multimodal investigation of different types of violations in lowly-constraining contexts.

5. Conclusions

Taken together, the results of these experiments constitute evidence that lexical prediction can occur across modalities and is modulated by contextual effects like sentential constraint in the same ways as unimodal prediction. This finding suggests a domain-independent processing, whereby substitutions across modalities are integrated into a singular multimodal utterance; lexical prediction appears to extend to these multimodal scenarios. This multimodal exploration of prediction meaningfully contributes to the ongoing theoretical discussions of the N400 and LFP. While the present research points strongly to big-picture "lexical" prediction of emoji, it does not bear on accounts of lower-level feature anticipation. As research continues on the potential anticipation of lower-level features of words, future work could probe whether lower-level features of emoji (i.e., visual characteristics) are anticipated in highly-constraining sentences. These insights may help to enlighten the unfolding discussion about feature anticipation (see e.g., Dikker et al., 2009; Eisenhauer et al., 2022; Gagl et al., 2020; Huang et al., 2022, and especially Nieuwland (2019) for a thorough critical review of many of the findings in this arena).

As this experiment utilized a between-rather than within-subject manipulation of the form of the sentence-final noun, it features more of these text-emoji substitutions than people naturally experience. While the multimodal parallel architecture emphasizes the ubiquity of multimodal interactions (e.g., co-speech gestures), the extent to which the frequency of a substitution affects the processing of its meaning is a worthy follow-up question. Future work should follow in the footsteps of Ganis et al. (1996) and investigate these multimodal substitutions rigorously as a within-subject manipulation to assess how substitution frequency affects meaning integration. It is possible that this within-subject manipulation may also provide insights as to the extension of the N400 effect and its appearance only to emoji stimuli in this experiment, though low-level visual-attentive differences between stimulus types may require clever alleviating.

This work contributes to the growing research on the language processing of emoji. Consistent with previous ERP work (Weissman and Tanner, 2018), we find that emoji presented in sentence contexts are processed in qualitatively similar ways to the words or tone that they replace. Irony marking and noun substitution are just two of the many and growing functions of emoji (Dainas and Herring, 2021), and additional work on the processing of these different graphic signs will continue to reveal how our linguistic system interacts with—or inherently reflects—multimodality. Emoji in multimodal contexts provide a natural setting in which to study fundamental processing across domains, and the findings presented here imply the integration of linguistic and pictorial processing. Continuing to study processing in these types of settings may provide valuable insights into the nature of language processing and domain-general cognition.

Data and materials

Data and analysis code are available at: https://osf.io/fygp5/?view_only=74bcbd7b2c2d4c7ebd76662d04a3a22d.

CRediT authorship contribution statement

Benjamin Weissman: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Neil Cohn:** Writing – review & editing, Supervision. **Darren Tanner:** Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

One author has advised in the creation of new emoji and in commercial use of emoji.

Data availability

Data and analysis code are available at: https://osf.io/fygp5/?view_only=74bcbd7b2c2d4c7ebd76662d04a3a22d

Acknowledgements

This work was supported by the National Science Foundation BCS-1431324. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Barach, E., Feldman, L.B., Sheridan, H., 2021. Are emojis processed like words?: eye movements reveal the time course of semantic processing for emojified text. Psychonomic Bull. Rev. https://doi.org/10.3758/s13423-020-01864-y.

Barber, H.A., Otten, L.J., Kousta, S.T., Vigliocco, G., 2013. Concreteness in word processing: ERP and behavioral effects in a lexical decision task. Brain Lang. 125 (1), 47–53. https://doi.org/10.1016/j.bandl.2013.01.005.

Bechtold, L., Bellebaum, C., Ghio, M., 2023. When a Sunny day gives you butterflies: an electrophysiological investigation of concreteness and context effects in semantic word processing. J. Cognit. Neurosci. 35 (2), 241–258. https://doi.org/10.1162/jocn_a_01942.

Bechtold, L., Ghio, M., Bellebaum, C., 2018. The effect of training-induced visual imageability on electrophysiological correlates of novel word processing. Biomedicines 6 (3). https://doi.org/10.3390/biomedicines6030075.

Beyersmann, E., Wegener, S., Kemp, N., 2023. That's good news ☹: semantic congruency effects in emoji processing. J. Media Psychol.: Theories, Methods, and Applications 35, 17–27. https://doi.org/10.1027/1864-1105/a000342.

Brehm, L., Alday, P.M., 2022. Contrast coding choices in a decade of mixed models. J. Mem. Lang. 125 (January), 104334 https://doi.org/10.1016/j.jml.2022.104334.

Brothers, T., Swaab, T., Traxler, M., 2015. Effects of prediction and contextual support on lexical processing: prediction takes precedence. Cognition 136 (Effects), 135–149. https://doi.org/10.1016/j.cognition.2014.10.017.

Brothers, T., Wlotko, E.W., Warnke, L., Kuperberg, G.R., 2020. Going the extra mile: effects of discourse context on two late positivities during language comprehension. Neurobiology of Language 1–58.

Brouwer, H., Crocker, M.W., Venhuizen, N.J., Hoeks, J.C.J., 2017. A neurocomputational model of the N400 and the P600 in language processing. Cognit. Sci. 41, 1318–1352. https://doi.org/10.1111/cogs.12461.

Carter, B.T., Foster, B., Muncy, N.M., Luke, S.G., 2019. Linguistic networks associated with lexical , semantic and syntactic predictability in reading: a fixation-related fMRI study. Neuroimage 189 (June 2018), 224–240.

Chen, L., Cichy, R.M., Kaiser, D., 2022. Semantic scene-object consistency modulates N300/400 EEG components, but does not automatically facilitate object representations. Cerebr. Cortex 32 (16), 3553–3567. https://doi.org/10.1093/cercor/bhab433.

Christofalos, A., Feldman, L., Sheridan, H., 2022. Semantic congruency facilitates memory for emojis. In: Proceedings of the the Fifth International Workshop on Emoji

Understanding and Applications in Social Media, pp. 63–68. https://doi.org/10.18653/v1/2022.emoji-1.7.
Clark, H.H., 1996. Using Language. Cambridge University Press, Cambridge, UK.
Coderre, E.L., O'Donnell, E., O'Rourke, E., Cohn, N., 2020. Predictability modulates neurocognitive semantic processing of non-verbal narratives. Sci. Rep. 10 (1), 1–11. https://doi.org/10.1038/s41598-020-66814-z.
Cohn, N., 2016. A multimodal parallel architecture : a cognitive framework for multimodal interactions. Cognition 146, 304–323. https://doi.org/10.1016/j.cognition.2015.10.007.
Cohn, N., 2021. A starring role for inference in the neurocognition of visual narratives. Cognitive Research: Principles and Implications 6 (1). https://doi.org/10.1186/s41235-021-00270-9.
Cohn, N., Engelen, J., Schilperoord, J., 2019. The grammar of emoji? Constraints on communicative pictorial sequencing. Cognitive Research: Principles and Implications 4 (1). https://doi.org/10.1186/s41235-019-0177-0.
Cohn, N., Foulsham, T., 2020. Zooming in on the cognitive neuroscience of visual narrative. Brain Cognit. 146 (October), 105634 https://doi.org/10.1016/j.bandc.2020.105634.
Cohn, N., Kutas, M., 2017. What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. Cognitive Research: Principles and Implications 2 (1). https://doi.org/10.1186/s41235-017-0064-5.
Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P.J., Kuperberg, G.R., 2012. (Pea)nuts and bolts of visual narrative: structure and meaning in sequential image comprehension. Cognit. Psychol. 65 (1), 1–38. https://doi.org/10.1016/j.cogpsych.2012.01.003.
Cohn, N., Roijackers, T., Schaap, R., Engelen, J., 2018. Are emoji a poor substitute for words? Sentence processing with emoji substitutions. In: Proceedings of the 40th Annual Conference of the Cognitive Science Society, pp. 1524–1529.
Cohn, N., Schilperoord, J., 2022. Remarks on multimodality: grammatical interactions in the parallel architecture. Frontiers in Artificial Intelligence 4 (January), 1–21. https://doi.org/10.3389/frai.2021.778060.
Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cognit. Neurosci. 6 (3), 256–266.
Coulson, S., Federmeier, K.D., Van Petten, C., Kutas, M., 2005. Right hemisphere sensitivity to word- and sentence-level context: evidence from event-related brain potentials. J. Exp. Psychol. Learn. Mem. Cognit. 31 (1), 129–147. https://doi.org/10.1037/0278-7393.31.1.129.
Cumming, G., 2014. The new statistics: why and how. Psychol. Sci. 25 (1), 7–29. https://doi.org/10.1177/0956797613504966.
Dainas, A., Herring, S.C., 2021. Interpreting emoji pragmatics. In: Xie, C., Yus, F., Haberland, H. (Eds.), Approaches to Internet Pragmatics, pp. 107–144. https://doi.org/10.1075/pbns.318.04dai.
DeLong, K.A., Quante, L., Kutas, M., 2014a. Predictability, plausibility, and two late ERP positivities during written sentence comprehension. Neuropsychologia 61 (1), 150–162. https://doi.org/10.1016/j.neuropsychologia.2014.06.016.
Delong, K.A., Troyer, M., Kutas, M., 2014b. Pre-processing in sentence comprehension: sensitivity to likely upcoming meaning and structure. Language & Linguistics Compass 8 (12), 631–645. https://doi.org/10.1111/lnc3.12093.Pre-processing.
DeLong, K.A., Urbach, T.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8 (8), 1117–1121. https://doi.org/10.1038/nn1504.
Dikker, S., Rabagliati, H., Pylkkänen, L., 2009. Sensitivity to syntax in visual cortex. Cognition 110 (3), 293–321. https://doi.org/10.1016/j.cognition.2008.09.008.
Dillon, B., Nevins, A., Austin, A.C., Phillips, C., 2012. Syntactic and semantic predictors of tense in Hindi: an ERP investigation. Lang. Cognit. Process. 27 (3), 313–344.
Dimitrova, D., Chu, M., Wang, L., Özyürek, A., Hagoort, P., 2016. Beat that word: how listeners integrate beat gesture and focus in multimodal speech discourse. J. Cognit. Neurosci. 28 (9), 1255–1269.
Eisenhauer, S., Gagl, B., Fiebach, C.J., 2022. Predictive pre-activation of orthographic and lexical-semantic representations facilitates visual word recognition. Psychophysiology 59 (3), 1–26. https://doi.org/10.1111/psyp.13970.
Federmeier, K.D., Kutas, M., 2001. Meaning and modality : influences of context , semantic memory organization , and perceptual predictability on picture processing. J. Exp. Psychol. Learn. Mem. Cognit. 27 (1), 202–224. https://doi.org/10.1037//027.
Federmeier, K.D., Mclennan, D.B., Ochoa, E.D.E., 2002. The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: an ERP study. Psychophysiology 39, 133–146.
Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M., 2007. Multiple effects of sentential constraint on word processing. Brain Res. 1146 (1), 75–84. https://doi.org/10.1016/j.brainres.2006.06.101.
Ferreira, F., Lowder, M.W., 2016. Prediction , information structure , and good-enough language processing. In: The Pschology of Learning and Motivation, vol. 65. https://doi.org/10.1016/bs.plm.2016.04.002.
Ferreira, F., Qiu, Z., 2021. Predicting syntactic structure. Brain Res. 1770 (August), 147632 https://doi.org/10.1016/j.brainres.2021.147632.
Filik, R., Turcan, A., Thompson, D., Harvey, N., Davies, H., Turner, A., 2015. Sarcasm and emoticons: comprehension and emotional impact. Q. J. Exp. Psychol. 218 (November), 1–17. https://doi.org/10.1080/17470218.2015.1106566.
Fleur, D.S., Flecken, M., Rommers, J., Nieuwland, M.S., 2020. Definitely saw it coming ? The dual nature of the pre-nominal prediction. Cognition 204, 104335. https://doi.org/10.1016/j.cognition.2020.104335.
Gagl, B., Sassenhagen, J., Haan, S., Gregorova, K., Richlan, F., Fiebach, C.J., 2020. An orthographic prediction error as the basis for efficient visual word recognition. Neuroimage 214 (December 2019), 116727. https://doi.org/10.1016/j.neuroimage.2020.116727.
Ganis, G., Kutas, M., Sereno, M.I., 1996. The search for "common sense": an electrophysiological study of the comprehension of words and pictures in reading. J. Cognit. Neurosci. 8 (2), 89–106.
Ge, J., Herring, S.C., 2018. Communicative functions of emoji sequences on Sina Weibo. Clin. Hemorheol. and Microcirc. 23 (11) https://doi.org/10.5210/fm.v23i11.9413.
Gibson, W., Huang, P., Yu, Q., 2018. Emoji and communicative action: the semiotics, sequence and gestural actions of 'face covering hand.'. Discourse, Context and Media 26, 91–99. https://doi.org/10.1016/j.dcm.2018.05.005.
Grosz, P.G., Kaiser, E., Pierini, F., 2023. Discourse anaphoricity vs. perspective sensitivity in emoji semantics. Glossa 8 (1), 1–35. https://doi.org/10.16995/glossa.9128.
He, Y., Luell, S., Muralikrishnan, R., Straube, B., Nagels, A., 2020. Gesture's body orientation modulates the N400 for visual sentences primed by gestures. Hum. Brain Mapp. 41 (17), 4901–4911. https://doi.org/10.1002/hbm.25166.
Heilbron, M., Armeni, K., Schoffelen, J.M., Hagoort, P., De Lange, F.P., 2022. A hierarchy of linguistic predictions during natural language comprehension. Proc. Natl. Acad. Sci. U.S.A. 119 (32), 1–12. https://doi.org/10.1073/pnas.2201968119.
Herring, S., Dainas, A., 2017. "Nice picture comment!" Graphicons in facebook comment threads. In: Proceedings of the Fiftieth Hawai'i International Conference on System Sciences (HICSS-50), pp. 2185–2194.
Hintz, F., Khoe, Y.H., Strauß, A., Psomakas, A.J.A., Holler, J., 2023. Electrophysiological evidence for the enhancement of gesture-speech integration by linguistic predictability during multimodal discourse comprehension. In: Cognitive, Affective and Behavioral Neuroscience.
Holcomb, P.J., MacPherson, W.B., 1994. Event-related brain potentials reflect semantic priming in an object decision task. Brain Cognit. 24 (2), 259–276.
Holler, J., Levinson, S.C., 2019. Multimodal Language processing in human communication. Trends Cognit. Sci. 23 (8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006.
Holtgraves, T., Robinson, C., 2020. Emoji can facilitate recognition of conveyed indirect meaning. PLoS One 15 (4), 1–13. https://doi.org/10.1371/journal.pone.0232361.
Huang, Z., Yang, S., Xue, L., Yang, H., Lv, Y., Zhao, J., 2022. Level of orthographic knowledge helps to reveal automatic predictions in visual word processing. Front. Neurosci. 15 (February), 1–8. https://doi.org/10.3389/fnins.2021.809574.
Jouen, A.L., Cazin, N., Hidot, S., Madden-Lombardi, C., Ventre-Dominey, J., Dominey, P. F., 2021. Common ERP responses to narrative incoherence in sentence and picture pair comprehension. Brain Cognit. 153, 105775 https://doi.org/10.1016/j.bandc.2021.105775.
Kretzschmar, F., Alday, P.M., 2023. Principles of statistical analyses: old and new tools. In: Grimaldi, M., Brattico, E., Shtyrov, Y. (Eds.), Language Electrified Principles, Methods, and Future Perspectives of Investigation. Springer, New York, pp. 123–160.
Kuperberg, G.R., Brothers, T., Wlotko, E.W., 2020. A Tale of Two Positivities and the N400: distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. J. Cognit. Neurosci. 32 (1), 12–35.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207 (4427), 203–205.
Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cognit. 11 (5), 539–550. https://doi.org/10.3758/BF03196991.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307 (5947), 161. https://doi.org/10.1038/307161a0.
Laszlo, S., Federmeier, K.D., 2011. The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. Psychophysiology 48 (2), 176–186. https://doi.org/10.1111/j.1469-8986.2010.01058.x.
Lau, E., Phillips, C., Poeppel, D., 2009. A cortical network for semantics:(de) constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933. https://doi.org/10.1038/nrn2532.
Lau, E., Polinsky, M., Clarke, N., Socolof, M., 2022. On the representation of morphosyntactic predictions: ERP evidence from Georgian. In: Makharoblidze, T. (Ed.), Issues in Kartvelian Studies. Vernon Press, p. 107.
Lau, E., Stroud, C., Plesch, S., Phillips, C., 2006. The role of structural prediction in rapid syntactic analysis. Brain Lang. 98, 74–88. https://doi.org/10.1016/j.bandl.2006.02.003.
Lee, C. lin, Federmeier, K.D., 2008. To watch, to see, and to differ: an event-related potential study of concreteness effects as a function of word class and lexical ambiguity. Brain Lang. 104 (2), 145–158. https://doi.org/10.1016/j.bandl.2007.06.002.
Leuthold, H., Kunkel, A., Mackenzie, I.G., Filik, R., 2015. Online processing of moral transgressions: ERP evidence for spontaneous evaluation. Soc. Cognit. Affect Neurosci. 10 (8), 1021–1029. https://doi.org/10.1093/scan/nsu151.
Li, L., Yang, Y., 2018. Pragmatic functions of emoji in internet-based communication—a corpus-based study. Asian-Pacific Journal of Second and Foreign Language Education 3 (1). https://doi.org/10.1186/s40862-018-0057-z.
Lopez-Calderon, J., Luck, S.J., 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front. Hum. Neurosci. 8, 213.
Luke, S.G., Christianson, K., 2015. Predicting inflectional morphology from context. Language, Cognition and Neuroscience 30 (6), 735–748. https://doi.org/10.1080/23273798.2015.1009918.
Luke, S.G., Christianson, K., 2016. Limits on lexical prediction during reading. Cognit. Psychol. 88, 22–60. https://doi.org/10.1016/j.cogpsych.2016.06.002.

Manfredi, M., Cohn, N., Kutas, M., 2017. Brain & Language when a hit sounds like a kiss : an electrophysiological exploration of semantic processing in visual narrative. Brain Lang. 169, 28–38. https://doi.org/10.1016/j.bandl.2017.02.001.

McNeill, D., 1992. Hand and Mind: what Gestures Reveal about Thought. University of Chicago Press, Chicago.

Meßmer, J.A., Bader, R., Mecklinger, A., 2021. Brain and Cognition the more you know: schema-congruency supports associative encoding of novel compound words . Evidence from event-related potentials. Brain Cognit. 155 (June) https://doi.org/10.1016/j.bandc.2021.105813.

Miller, H., Kluver, D., Thebault-Spieker, J., Terveen, L., Hecht, B., 2017. Understanding emoji ambiguity in context: the role of text in emoji-related miscommunication. In: Proceedings of the Eleventh International AAAI Conference on Web and Social Media, pp. 152–161. Retrieved from. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15703.

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., Hecht, B., 2016. "Blissfully happy" or "ready to fight": varying interpretations of emoji. In: *International AAAI Conference On Web and Social Media,* (Icwsm), pp. 259–268. https://doi.org/10.1089/cyber.2011.0179.

Morett, L.M., Landi, N., Irwin, J., McPartland, J.C., 2020. N400 amplitude, latency, and variability reflect temporal integration of beat gesture and pitch accent during language processing. Brain Res. 1747 (March), 147059 https://doi.org/10.1016/j.brainres.2020.147059.

Na'aman, N., Provenza, H., Montoya, O., 2017. Varying linguistic purposes of emoji in (Twitter) context. Proceedings of ACL 2017. Student Research Workshop, pp. 136–141.

Nieuwland, M.S., 2019. Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. Neurosci. Biobehav. Rev. 96 (May 2018), 367–400. https://doi.org/10.1016/j.neubiorev.2018.11.019.

Nieuwland, M.S., Arkhipova, Y., Rodr\'\iguez-Gómez, P., 2020. Anticipating words during spoken discourse comprehension: a large-scale, pre-registered replication study using brain potentials. Cortex 133, 1–36.

Nieuwland, M.S., Barr, D.J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D.I., et al., 2019. Dissociable effects of prediction and integration during language comprehension: evidence from a largescale study using brain potentials. Phil. Trans. Biol. Sci. 375 (1791) https://doi.org/10.1098/rstb.2018.0522.

Nieuwland, M.S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., et al., 2018. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. Elife 7, 1–24. https://doi.org/10.7554/eLife.33468.

Nigam, A., Hoffman, J.E., Simons, R.F., 1992. N400 to semantically anomalous pictures and words. J. Cognit. Neurosci. 4 (1), 15–22. https://doi.org/10.1162/jocn.1992.4.1.15.

Nour Eddine, S., Brothers, T., Kuperberg, G.R., 2022. The N400 in silico: a review of computational models. In: Federmeier, K. (Ed.), Psychology of Learning and Motivation - Advances in Research and Theory, vol. 76. https://doi.org/10.1016/bs.plm.2022.03.005.

Paggio, P., Tse, A.P.P., 2022. Are emoji processed like words? An eye-tracking study. Cognit. Sci. 46 (2) https://doi.org/10.1111/cogs.13099.

Payne, B.R., Lee, C.-L., Federmeier, K.D., 2015. Revisiting the incremental effects of context on word processing: evidence from single-word event-related brain potentials. Psychophysiology 52 (11), 1456–1469.

Pérez, A., Schmidt, E., Kourtzi, Z., Tsimpli, I., 2020. Multimodal semantic revision during inferential processing: the role of inhibitory control in text and picture comprehension. Neuropsychologia 138 (March 2019). https://doi.org/10.1016/j.neuropsychologia.2019.107313.

Pfeifer, V.A., Armstrong, E.L., Lai, V.T., 2022. Do all facial emojis communicate emotion? The impact of facial emojis on perceived sender emotion and text processing. Comput. Hum. Behav. 126, 107016.

Poulton, V.R., Nieuwland, M.S., 2022. Can you hear What's coming? Failure to replicate ERP evidence for phonological prediction. Neurobiology of Language 3 (4), 556–574. https://doi.org/10.1162/nol_a_00078.

Pratha, N.K., Avunjian, N., Cohn, N., 2016. Pow , punch , Pika , and Chu : the structure of sound effects in Genres of American comics and Japanese manga. Multimodal Communication 5 (2), 93–109. https://doi.org/10.1515/mc-2016-0017.

Quante, L., Bölte, J., Zwitserlood, P., 2018. Dissociating predictability, plausibility and possibility of sentence continuations in reading: evidence from late-positivity ERPs. PeerJ 1–24. https://doi.org/10.7717/peerj.5717.

Rabovsky, M., Hansen, S.S., McClelland, J.L., 2018. Modelling the N400 brain potential as change in a probabilistic representation of meaning. Nat. Human Behav. 2 (9), 693–705. https://doi.org/10.1038/s41562-018-0406-4.

Robus, C.M., Hand, C.J., Filik, R., Pitchford, M., 2020. Investigating effects of emoji on neutral narrative text: evidence from eye movements and perceived emotional valence. Comput. Hum. Behav. 109 (April), 106361 https://doi.org/10.1016/j.chb.2020.106361.

Rodríguez-Gómez, P., Romero-Ferreiro, V., Pozo, M.A., Hinojosa, J.A., Moreno, E.M., 2020. Facing stereotypes: ERP responses to male and female faces after gender-stereotyped statements. Soc. Cognit. Affect Neurosci. 15 (9), 928–940. https://doi.org/10.1093/scan/nsaa117.

Romero-Rivas, C., Corey, J.D., Garcia, X., Thierry, G., Martin, C.D., Costa, A., 2017. World knowledge and novel information integration during L2 speech comprehension. Biling. Lang. Cognit. 20 (3), 576–587.

Schad, D.J., Vasishth, S., Hohenstein, S., Kliegl, R., 2020. How to capitalize on a priori contrasts in linear (mixed) models: a tutorial. J. Mem. Lang. 110 (July 2019), 104038 https://doi.org/10.1016/j.jml.2019.104038.

Scheffler, T., Brandt, L., Fuente, M. de la, Nenchev, I., 2022. The processing of emoji-word substitutions: a self-paced-reading study. Comput. Hum. Behav. 127 (August 2021), 107076 https://doi.org/10.1016/j.chb.2021.107076.

Schlenker, P., 2019. Gestural semantics: replicating the typology of linguistic inferences with pro- and post-speech gestures. Nat. Lang. Ling. Theor. 37 (2), 735–784. https://doi.org/10.1007/s11049-018-9414-3.

Tang, M., Chen, B., Zhao, X., Zhao, L., 2020. Processing network emojis in Chinese sentence context: an ERP study. Neurosci. Lett. 722 (November 2019), 134815 https://doi.org/10.1016/j.neulet.2020.134815.

Tang, M., Zhao, X., Chen, B., Zhao, L., 2021. EEG theta responses induced by emoji semantic violations. Sci. Rep. 1-9, 123456789 https://doi.org/10.1038/s41598-021-89528-2.

Tanner, D., 2019. Robust neurocognitive individual differences in grammatical agreement processing: a latent variable approach. Cortex 111, 210–237. https://doi.org/10.1016/j.cortex.2018.10.011.

Taylor, W.L., 1953. "Cloze procedure": a new tool for measuring readability. Journalism Bulletin 30 (4), 415–433. https://doi.org/10.1177/107769905303000401.

Thornhill, D.E., Van Petten, C., 2012. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. Int. J. Psychophysiol. 83 (3), 382–392. https://doi.org/10.1016/j.ijpsycho.2011.12.007.

Tieu, L., Schlenker, P., Chemla, E., 2019. Linguistic inferences without words. Proc. Natl. Acad. Sci. U.S.A. 116 (20), 9796–9801. https://doi.org/10.1073/pnas.1821018116.

Tigwell, G.W., Flatla, D.R., 2016. Oh that ' s what you meant !": reducing emoji misunderstanding. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct, pp. 859–866. https://doi.org/10.1145/2957265.2961844.

Van Den Brink, D., Brown, C.M., Hagoort, P., 2001. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. J. Cognit. Neurosci. 13 (7), 967–985.

Van Petten, C., Luka, B.J., 2012. Prediction during language comprehension: benefits , costs , and ERP components. Int. J. Psychophysiol. 83, 176–190.

Vega-Mendoza, M., Pickering, M.J., Nieuwland, M.S., 2021. Neuropsychologia Concurrent use of animacy and event-knowledge during comprehension: evidence from event-related potentials. Neuropsychologia 152 (November 2019), 107724. https://doi.org/10.1016/j.neuropsychologia.2020.107724.

Võ, M.L.H., Wolfe, J.M., 2013. Differential electrophysiological signatures of semantic and syntactic scene processing. Psychol. Sci. 24 (9), 1816–1823. https://doi.org/10.1177/0956797613476955.

Wang, T., Xu, X., 2022. Brain and Language Better I than He : personal perspective modulates counterfactual processing. Brain Lang. 228 (March), 105105 https://doi.org/10.1016/j.bandl.2022.105105.

Weissman, B., Engelen, J., Baas, E., Cohn, N., 2023. The lexicon of emoji? Conventionality modulates processing of emoji. Cognit. Sci. 47 (4).

Weissman, B., Tanner, D., 2018. A strong wink between verbal and emoji-based irony: how the brain processes ironic emojis during language comprehension. PLoS One 13 (8). https://doi.org/10.1371/journal.pone.0201727.

West, W.C., Holcomb, P.J., 2000. Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. J. Cognit. Neurosci. 12 (6), 1024–1037. https://doi.org/10.1162/08989290051137558.

Wlotko, E.W., Federmeier, K.D., 2012. So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of messagelevel meaning. Neuroimage 62 (1), 356–366. https://doi.org/10.1016/j.neuroimage.2012.04.054.So.

Yang, X., Liu, M., 2021. The pragmatics of text-emoji co-occurrences on Chinese social media. Pragmatics 31 (1), 144–172. https://doi.org/10.1075/prag.18018.yan.

Yus, F., 2021. Smartphone communication. In: Interactions in the App Ecosystem. Routledge, London.

Zhang, Y., Frassinelli, D., Tuomainen, J., Skipper, J.I., Vigliocco, G., 2021. More than words: word predictability, prosody, gesture and mouth movements in natural language comprehension. Proc. Biol. Sci. (1955), 288. https://doi.org/10.1098/rspb.2021.0500.